Breast cancer detection using Deep Learning

Nguyen Trung Hieu

Hanoi University of Science and Technology

BK.AI – Computer vision Research group

1) INTRODUCTION

Breast cancer is the type of cancer that originates in your breast tissue. It occurs when breast cells mutate (change) and grow out of control, creating a mass of tissue (tumor). Like other cancer, breast cancer can invade and grow into the tissue surrounding your breast. It can also travel to other parts of your body and form new tumors.

Breast cancer is the second most cancer in women, after skin cancer. According to GLOBOCAN, there are 21555 new cases of breast cancer in Vietnam in 2020, accounting for 11.8% of new cases in 2020, only after liver cancer and lung cancer. And in Vietnam, most of the cases are detected at later stages due to lack of awareness of breast cancer, which leads to difficult and expensive treatments. Detection of breast cancer in the early stages is crucial for proper treatment and reducing the mortality rate. Breast cancer is mostly detected through mammography, but the accuracy of the result depends heavily on the expertise of the radiologist, which might lead to false positives.

In this project, I propose using the Faster R-CNN network with Feature pyramid network for the detection of suspicious lesions in mammograms to support the radiologist in detecting breast cancer. Note that not all abnormalities in the breast are malignant, most of them can be benign.

The remainder of the report is organized as follows: section 2 is detailed architecture of the network used, section 3 is the details of the experiment and results and the references.

2) DEEP LEARNING MODEL: FASTER R-CNN WITH FEATURE PYRAMID NETWORK

2.1 FASTER R-CNN

2.1.1 REGION PROPOSAL NETWORK (RPN):

This is a fully convolutional network that takes an image of any size and outputs a set of regions proposed, each with an objectiveness score. The RPN consists of 2 parts: a shareable backbone network and a head module.

We pass the image through the backbone to obtain a feature map with lower resolution. Then we slide the head module through the feature map to generate the region proposal. The head module takes an n x n spatial window as input, mapped to a 256-d vector, and feeds to 2 layers: a box regression layer and a box classification layer.

To achieve translation and invariant, we use the anchor box idea. At each location of the feature map, we consider a set of k anchors with different sizes and ratios instead of 1 box. The classification layer will have 2k outputs and the regression layer has 4k outputs with k anchors.

To train the RPN, we first mark each anchor with either positive or negative. A positive anchor is either the anchor that has IoU highest with the ground-truth box or the anchor that IoU of that anchor with any ground-truth box is higher than 0.7. A negative anchor is the anchor has IoU less than 0.3 for all ground-truth boxes. Any boxes that are not positive or negative are ignored during training.

We train the RPN with 2 losses. The first one is the log loss over 2 classes: object or not an object, on all negative and positive boxes. The second losses is the bounding box regression for positive boxes only, we use the Robust loss function define as:

$$R(x) = \begin{cases} 0.5x^2 \; if \;\; |x| < 1 \\ |x| - 0.5 \; otherwise \end{cases}$$

The loss function for one image is:

$$L(\{p_i\}, \{t_i\}) = \frac{1}{N_{cls}} \sum L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum p_i^* R(t_i - t_i^*)$$

where i is the index of the image, p_i and t_i is the predicted class and coordinate of the bounding box, p_i^* and t_i^* are the ground-truth class and coordinate of the bounding box, L_{cls} is the log loss over 2 class object and the background. N_{cls} and N_{reg} are normalization term, and λ is the weighted parameter between 2 loss.

2.1.2 FAST R-CNN MODEL:

After getting a set of proposed regions from RPN, we use these regions for prediction. First, we still pass the image through the shareable backbone used in RPN to obtain a feature map. And for each region proposed a RoI pooling layer is applied to extract a fixed size feature vector from the feature map. And each feature vector is passed through a small network to 2 output layers: a classification layer for (K + 1) object (K object and the background) and a box regression layer.

A RoI pooling layer takes a region proposed and mapped it on the feature map, and then divides the region on the feature map into a 7x7 grid window. After that, applying max pooling on each subgrid, we obtain a 7x7 feature vector.

To train the Fast R-CNN model, we also train with 2 losses like the RPN, we just need to replace the log loss over 2 classes with log loss over (K + 1) classes.

2.1.3 TRAIN THE FASTER R-CNN MODEL:

We have seen the Faster R-CNN model as 2 independent modules with a shareable backbone. To utilize parallel computing and GPU power, we need to train the whole network end-to-end as an unify network. We use alternating training strategy: we first train the RPN, then use the proposals to train Fast R-CNN. The network is fine-tuned by training Fast R-CNN is used to initialize RPN, then the process is iterated.

2.2 FEATURE PYRAMID NETWORK:

2.2.1 DEFINITION:

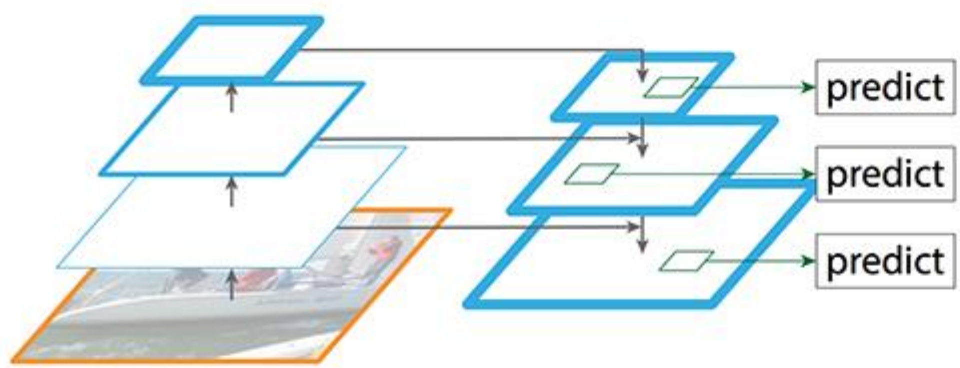


Fig 1: Feature pyramid network architecture. Source

Feature pyramid network (FPN) is a feature extractor at multiple levels and fully convolutional. FPN consists of a bottom-up pathway, a top-down pathway and lateral connection.

Bottom-up pathway is the feed forward in the convolutional neural network. There are layers in the network that output the feature map with the same size, we put these layers in a module. For our pyramid, we define one pyramid level for each module, the last layer of each module is the feature map of that level in the pyramid.

Top-down pathway and lateral connection: from the feature at the top of the pyramid we apply 1x1 convolution to reduce channels and upsampling to match the size of the feature at the lower level. We merge it with the feature at the lower level feature to obtain the final feature map at that level

2.2.2 FEATURE PYRAMID NETWORK WITH RPN

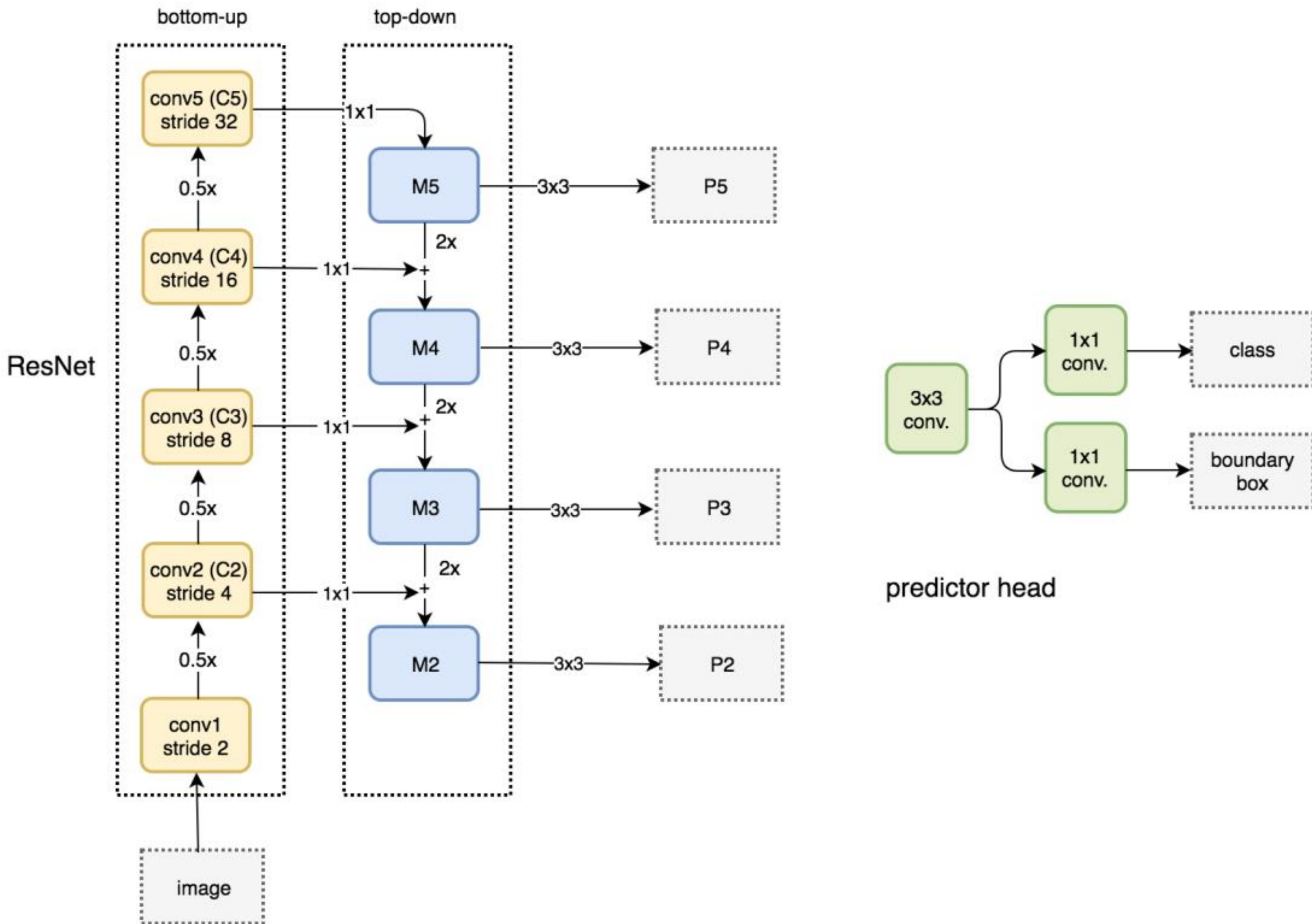


Fig 2: Feature Pyramide network with RPN. Source

In the original RPN, we apply multiple anchors at each location to cover various types of objects at different sizes and scales. We apply the FPN in the RPN to generate multiple feature maps of the image instead of a single feature map. Then we apply a single anchor box in each feature map of the pyramid instead of multiple anchor boxes at each point of a single feature map.

2.2.3 FEATURE PYRAMID NETWORK WITH FAST R-CNN MODEL

In the original Fast R-CNN we apply RoI pooling for a region proposed on a single feature map. But with FPN, multiple feature maps are generated, so we need to choose the feature map to apply the RoI on. We find the pyramid level by the formula:

$$k = \lfloor k_0 + \log_2 \sqrt{wh}/224 \rfloor$$

where k_0 is the top level of the pyramid, 224 is the ImageNet image size

3) EXPERIMENTAL RESULTS:

3.1 DATASETS:

In this project we use the mini MIAS dataset. The mini MIAS dataset contains 322 mammograms from 161 patients (left breast and right breast). The dataset also contains radiologist ground-truth abnormality if it may present. The database has been reduced to a 200 micron pixel edge and padded/clipped so that all the images are 1024x1024.

3.2 NETWORK SPECIFICATIONS:

In this project we use Faster R-CNN with Feature Pyramid network to detect abnormality in the mammograms. We won't classify the abnormality is benign or malignant.

The dataset ground truth comes in the form circle with coordinate (x, y) and radius R of the circle. We transform (x, y, R) to rectangle with (x_min, x_max, y_min, y_max) as (x - R, x + R, y - R, y + R).

Initially, the dataset has 115 mammograms with abnormality (positive images) and 207 mammograms with no abnormality (negative images). Through preprocessing, we found that there are 4 positive images with no coordinate data and 2 positive images where the abnormality is outside the breast, so we have only 109 positive images for training and testing. We split positive images and negative images into training set, validation set, test set with ratio 6/2/2.

We also apply online augmentation on the dataset, flip the image horizontally.

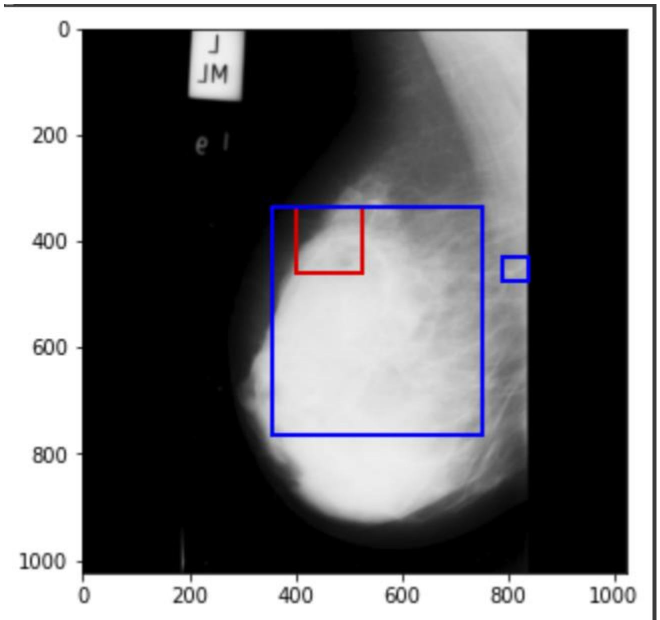
In the Faster R-CNN model, we use Resnet50 as the backbone network, with anchor sizes (32, 64, 128, 256, 512) and ratios (0.5, 1, 2). The backbone is pretrained on Coco 2017 Dataset, and freeze the first 2 layers while training the network.

We train the network with the SGD optimizer, learning rate is 0.005, momentum is 0.9 and L2 loss with $\lambda = 0.0005$. We schedule the learning rate decrease 10 times after 3 epochs. We train the model for 9 epochs.

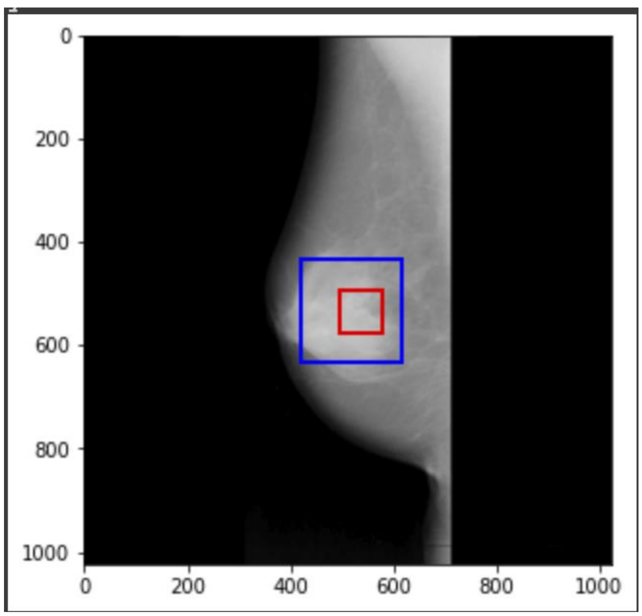
3.3 RESULTS

After we evaluate the model on the test set, we get the following results: the mean Average precision (mAP) average over IoU from 0.5 to 0.95 is mAP = 13.4%, the Average precision with IoU = 0.5 is AP = 26.4%, the Average precision with IoU = 0.75 is 13.8%.

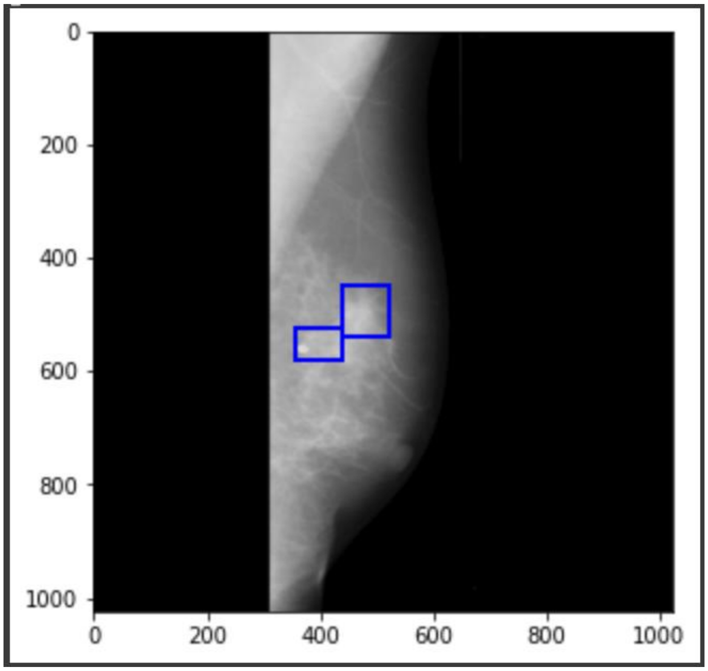
Compared to the same model architecture trained on Coco dataset 2017 and achieving mAP = 37.8%, this is a bad result. This result mainly comes from the problem with the mini MIAS dataset. This is a small dataset with 322 mammograms, even though we use fine-tuning to train the model, poor results are unavoidable. Also, mini MIAS dataset is a very imbalanced dataset, where in 322 mammograms there are 207 mammograms that have no abnormality, and only 115 mammograms have abnormality. This imbalance has caused great challenges for the model to learn the dataset, since it learns too many background classes, and little object classes lead to poor results. Finally, noisy data in the dataset is also the problem, where the coordinate of some mammograms is not accurate.



a) Mammogram 1



b) Mammogram 2



c) Mammogram 3

Fig 3: output of the model. Red boxes are ground-truth boxes, blue boxes are prediction boxes

We visualize some of the results on the test set. 2 main problems with our model is that its prediction on small abnormalities is not very accurate, and a high number of false positives. On the first mammogram and the second mammogram, we see that the ground-truth box in red is small, where the predicted box in blue is much larger, though it does cover the ground-truth box. This is not a very big problem, because the radiologist will still know how to find the abnormality with his or her expertise. And on the third mammogram, though there is no abnormality on it, the model still outputs a bounding box. This is a problem because these kinds of predictions will lead to a high chance of false positives and misdiagnosis.

3.4 PROPOSALS FOR ENHANCEMENTS:

To improve the models, i have some proposals:
- Apply more heavy augmentations: horizontal and vertical flip, add noise, cropping, …
- Pretrain the model on a larger medical dataset.
- Use another pretrained backbone network: Larger backbone ResNet 101 with FPN, ResNet 152 with FPN, Inception Resnet, …
- Ensembling multiple object detection models instead of a single one and average the results.

References

[1] Girshick R, Donahue J, Darrell T, Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation. In: 2014 IEEE Conference on Computer Vision and Pattern Recognition. IEEE; 2014.

[2] Girshick R. Fast R-CNN. 2015; Available from: http://dx.doi.org/10.48550/ARXIV.1504.08083

[3] Ren S, He K, Girshick R, Sun J. Faster R-CNN: Towards real-time object detection with region proposal networks [Internet]. arXiv [cs.CV]. 2015. Available from: http://arxiv.org/abs/1506.01497

[4] Lin T-Y, Dollár P, Girshick R, He K, Hariharan B, Belongie S. Feature pyramid networks for object detection. 2016; Available from: http://dx.doi.org/10.48550/ARXIV.1612.03144

[5] Tsochatzidis L, Costaridou L, Pratikakis I. Deep learning for breast cancer diagnosis from mammograms-A comparative study. J Imaging [Internet]. 2019;5(3):37. Available from: http://dx.doi.org/10.3390/jimaging5030037

[6] Hui J. Understanding Feature Pyramid Networks for object detection (FPN) [Internet]. Medium. 2018. Available from: https://jonathan-hui.medium.com/understanding-feature-pyramid-networks-for-object-detection-fpn-45b227b9106c

[7] What is breast cancer? [Internet]. Cancer.org. [cited 2022 Mar 23]. Available from: https://www.cancer.org/cancer/breast-cancer/about/what-is-breast-cancer.html